CHAPTER 14

Identification of Feigned Mental Retardation

Tara L. Victor
Kyle Brauer Boone

Individuals may choose to feign or exaggerate intellectual deficits for many reasons. External incentives might include receiving extra time on tests, monetary compensation, or evading occupational or military obligations, but in some cases the stakes are much higher. In the forensic arena, a defendant's competency to stand trial (CST) is, in part, determined by his or her itellectual ability. In fact, CST has been one of the most studied areas of mental health law since the case of *Jackson v. Indiana* in 1972 in which a mentally retarded defendant (who was also unable to speak, hear, read, or write) was charged with two robberies but found incompetent. Typically when this happens (e.g., as in the case of psychotic disturbance), the defendant is institutionalized with the goal of restoring competency and eventually trying the case in court; however, in this particular instance (i.e., a case of congenital mental deficiency), it was thought that competency would never be regained. As it was considered unconstitutional to institutionalize the defendant indefinitely, the court decided the defendant must either be released or committed in accordance with the general civil commitment provisions (see Schlesinger, 2003). Thus, the incentive for successful feigning of intellectual impairment is clear; it has the potential to allow a defendant to avoid the consequences of a major criminal offense, and in fact, the base rate for feigned cognitive impairment in a CST sample has been estimated at 13–17% (Heinze & Purisch, 2001; Frederick, 2000a, 2000b).

Compounding the need to accurately assess effort in forensic neuropsychological assessment is the fact that, as of 2002, successful feigning of intellectual and functional impairment may allow criminals to avoid the death penalty. More specifically, in the recent Supreme Court case of *Atkins v. Virginia* (2002), the Court determined in a 6–3 decision that executing the mentally retarded was "cruel and unusual" punishment and therefore prohibited by the Eighth Amendment (see Keyes, Edwards, & Perske, 1997; for an excellent description of case details, see Graue et al., in press). Given that defendants must have the capacity to assist their attorneys in their defense (American Bar Association, 1984) and there is evidence to suggest that MR individuals are more susceptible to making false confessions (Gudjonsson, 2002, as cited in Johnstone & Cooke, 2003), as well as inaccurate statements, during testimony (Everington & Fulero, 1999), a number of states had already outlawed this practice. The federal court majority, however, reportedly wanted to create greater consistency with their decision. It left the definition of mental retardation (MR) up to the states, most of which define a person with MR as having an IQ of 70 or lower and having two or more adaptive deficiencies, a topic that is addressed in more detail later.

Earlier reports suggested that approximately 16% of murder/insanity defendants were mentally retarded (Lanzkron, 1963). More recent estimates indicate that between 4% and 10% of the jail population in the United States comprises individuals with MR (Davis, 2000), and approximately 11% of inmates in maximum security and on death row meet criteria for the diagnosis (Everington & Keyes, 1999). While some writers have argued that inmates will want to mask their intellectual deficits given the assumed stigma associated with such a label, as well as the possibility that their disability might make them more vulnerable once they return to the general prison population (Dolan, 2005; Hawthorne, 2005), it is the opinion of others (and we agree) that the Supreme Court decision in 2002 will likely increase the number of assessments of MR in forensic settings and, consequently, the need for accurate assessment of malingering in such cases (Brodsky & Galloway, 2003). In fact, empirical research shows that criminal defendants facing serious charges are more likely to feign than those facing less serious charges (Weinborn, Orr, Woods, Conover, & Feix, 2003). This is of particular concern given that the MR population has been virtually ignored in the malingering literature and mentally retarded subjects are frequently excluded from validation studies of effort indices.

In sum, it is important to address the issue of noncredible cognitive test performance in the assessment of MR, as the incentive for feigning (at least in the forensic arena) is now literally a matter of life or death. However, the validity of using standard measures of effort with this population is unclear. With this in mind, the goals of this chapter are to (1) clarify the definition of MR and describe its general clinical characteristics; (2) identify the most common causes of MR and their specific neuropsychological correlates to

guide predictions about this population's ability to pass standard effort test cutoffs; (3) present the results of a recent survey sent out to diplomates in forensic psychology regarding current practice in this area; (4) review the extant empirical literature with respect to the validity of such practice and present preliminary data from our lab on the use of common effort indicators in a low IQ sample; and, finally, (5) integrate this information in such a way as to stimulate and guide the emergence of future research. Each of these areas is addressed in turn.

DEFINITION AND CLINICAL CHARACTERISTICS OF MR

revalence

Based on the normal distribution, 2–3% of the population (or about 7 million people in the United States) would be expected to have an IQ less than 70 (Pulsifier, 1996). This is consistent with epidemiological studies estimating approximately 6.2 to 7.5 million individuals are mentally retarded in the United States, 26,500 to 32,500 of which are in prison or residential facilities (Davis, 2000). Epidemiological studies based solely on the criteria of IQ less than 70 indicate prevalence rates ranging from 0.8 to 1.2% with approximately 3–6 per 1,000 in mild range, 2 per 1,000 in the moderate range, 1.3 per 1,000 in the severe range, and finally, 0.4 per 1,000 who meet criteria for profound MR (McLaren & Bryson, 1987; Lipkin, 1991). As one author points out, however, when adaptive criteria are included in the definition, the prevalence slightly lowers (Pulsifier, 1996). Estimates are also lowered by premature mortality (especially in profound MR; Pulsifier, 1996).

finitions

The aforementioned caveats surrounding estimations of prevalence highlight the fact that there are varying ways in which MR has been defined. The fourth edition, text revision of *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV-TR; American Psychiatric Association, 2000) defines MR as the presence of "subaverage intelligence" defined as an IQ less than or equal to 70 on the basis of formal testing or clinical judgment (i.e., in the case of an infant or an untestable subject) *and* "concurrent . . . impairments in present adaptive functioning" found in at least 2 of 11 areas of functional independence (i.e., communication, self-care, home living, social/interpersonal skills, use of community resources, self-direction, functional academic skills, work, leisure, health and safety) with onset before age 18 (p. 46). Here, adaptive functioning is defined as "the person's effectiveness in meeting the standards expected for his or her age and by his or her cultural group" (p. 46). DSM-IV-TR further classifies MR individuals by the IQ range within which they fall, the vast majority of cases falling into the category of mild MR:

Level of functioning	IQ range	% of all MR cases
Mild	50–55 to approximately 70	85%
Moderate	35–40 to 50–55	10%
Severe	20–25 to 35–40	3–4%
Profound	below 20–25	1–2%

The American Association on Mental Retardation (AAMR; 1992, 2002) has actually changed the definition of MR 10 times since 1908, in part spurred by a California class action lawsuit that found schools had placed a disproportionate number of black children in special programming for MR. For example, in 1973, an IQ of less than 85 was thought to be indicative of MR; however, this was eventually lowered to 70 and is now at 75 (allowing for a margin of error in testing). Notably, it has been suggested that this higher cutoff doubles the number of people eligible for diagnosis (Mash & Wolfe, 2002). Also, in contrast to DSM-IV-TR's focus on degree of impairment (i.e., mild, moderate, severe, and profound), the AAMR describes an individual's level of functioning in terms of his or her use of, and need for, external supports (i.e., intermittent, limited, extensive, or pervasive; Loveland & Tunali-Kotoski, 1998; Luckasson et al., 1992; also see www.aamr.org). Thus, the focus of this system of classification is on the individual's ability to function and required level of support as opposed to the degree of psychometric deficit found upon neuropsychological examination. Indeed, since the late 1970s and early 1980s, the importance of considering adaptive behavior (and not just IQ) in making the diagnosis has grown in awareness (DeVault & Long, 1988; Huberty, Koller, & Ten Brink, 1980). Authors have speculated several reasons for this, including (1) the identified need to evaluate such abilities in context (i.e., in the interaction between a person and their environment; Hawkins & Cooper, 1990); (2) an attempt to promote nondiscriminatory testing procedures, as low test scores are overrepresented in some minority groups (DeVault & Long, 1988; Everington & Keyes, 1999); (3) the ease with which identified needs for assistance can be translated into specific treatment goals (Mash & Wolfe, 2002); and finally (4) the general dissatisfaction with our existing practice of focusing solely on IQ and assumed disability versus actual need (see Bruininks, Thurlow, & Gilman, 1987, for a review of the evolution of including adaptive criteria). The last reason appears most strongly related to the fact that many mild MR individuals are capable of supporting themselves and living independently in the community (American Psychiatric Association, 2000); thus, psychometric deficit in these cases does not translate into actual adaptive deficit.

As defined by the American Association on Mental Deficiency (AAMD), deficits in adaptive functioning refer to "limitations in an individual's effectiveness in meeting the standards of maturation, learning, personal independence and/or social responsibility that are expected for his or her age

level . . ." (Grossman, 1983, p. 11). Elsewhere in the literature, adaptive functioning has also been defined as "the extent to which an individual takes care of personal needs, exhibits social competence, and refrains from exhibiting problem behaviors . . ." (Bruininks, McGrew, & Maruyama, 1988, p. 266), abilities which undoubtedly depend on age (Bruininks, Thurlow, & Gilman, 1987):

Age level	Expected adaptive skills
Infancy/early childhood	Sensorimotor skills, communication skills, self-help skills and socialization
Childhood/early adolescence	Application of academic skills to daily life, application of reasoning and judgment to environmental situations, development of group and interpersonal social skills
Late adolescence/adulthood	Vocational adjustment and performance as well as social adjustment in the community

Commonly used measures of adaptive functioning include the Vineland Adaptive Behavior Scales–II (VABS-II; Sparrow, Balla, & Cicchetti, 1984), the AAMD's Adaptive Behavior Scale (AAMD-ABS; Nihira, Foster, Shellhaas, & Leland, 1974), and the Adaptive Behavior Scale (ABS; Fogelman, 1975; Nihira, Leland, & Lambert, 1993), and literature focused on the assessment of adaptive behavior specifically in defendants charged with capital murder is emerging (Stevens & Randall, 2006).

General Clinical Characteristics

Mental retardation occurs more often in men than women by a ratio of approximately 1.6 to 1.0 (see Pulsifier, 1996), and can occur through a number of varied etiologies, including "infection and intoxication, trauma or physical agents, disorders of metabolism or nutrition, gross prenatal brain disease, unknown prenatal factors, chromosomal abnormalities, gestational disorders, environmental factors, secondary to psychiatric conditions, and other conditions" (McCaffrey & Isaac, 1985, p. 63). Early research did not differentiate between groups of varying etiology and instead simply investigated the characteristics of samples of individuals with similar IQ (e.g., Roszkowski & Snelbecker, 1981). This early research hypothesized that individuals with MR had most difficulty with what was referred to at the time as short-term memory (Ellis, 1963) but what today would be referred to as attention and speed of information processing. For example, Sprague and Quay (1966) factor-analyzed the performance of a heterogeneous group of subjects with MR on the Wechsler Adult Intelligence Scale (WAIS) and identified four factors, including what the authors referred to as a trace factor (arithmetic, digit span and digit symbol) on which the subjects with MR performed the worst. More recently, MacKenzie and Hulme (1987) found significantly lower than expected digit spans (relative to mental age), and significantly slowed development of digit span as mental age increased over time, in a group with severe MR relative to controls. In addition, adults with MR were found to demonstrate slower mental speed measured through analysis of eye gaze, eye movement, and orientation as compared to controls (Nettelbeck, 1986).

Other research using similar samples and research methodology demonstrated deficits in other cognitive domains, such as executive/frontal dysfunction, including deficits in planning and regulation of simple behavioral tasks (McCaffrey & Isaac, 1985). This finding is of particular note, as adaptive functions are thought to be most strongly related to the cognitive domain of executive function in an MR population (Edgin, 2003). Limitations of this work, however, include not only the fact that MR etiology went unaccounted for but also that this particular study included an extremely small sample size ($n = 10$). Other cognitive profiles emerging from this literature include greater difficulty with arithmetic as compared to reading (see Pulsifier, 1996) and a marked discrepancy between Verbal and Performance IQs, the former of which was found to be significantly lower than the latter (Calvert & Crozier, 1978). Taken together, the above-mentioned findings led researchers to believe the primary neuropsychological difficulty in MR is in sequential information processing, which involves primarily attention, processing speed, and planning abilities. This type of processing would be in contrast to what is known as simultaneous processing, relying more on perceptual integration and spatial skills (Pulsifier, 1996).

Previous research has also documented deficits in motor speed, grip strength, and sensory function within this population. For example, subjects with mild and moderate MR were found to exhibit slower physical speed than controls in one study (Lin & Zhang, 2002) and slower physical speed was highly related to IQ in 245 mentally retarded individuals assigned to groups on the basis of IQ (Black & Davis, 1966). Further, decreased grip strength was found in children with mild to severe MR when compared to a non-mentally retarded group (Iki & Kusano, 1986). Finally, both visual and hearing impairments have been documented in MR samples (Pulsifier, 1996).

Although these general profiles have emerged, as mentioned earlier, the etiology of MR is quite diverse. More recent research has investigated the possibility that the neuropsychological profiles of groups of varying etiology may differ as well. Indeed, emerging from this literature is evidence to suggest the existence of different cognitive profiles and neuroanatomical deficits associated with the different causes of MR.

CAUSES OF MR AND THEIR NEUROPSYCHOLOGICAL CORRELATES

The patterns of cognitive and adaptive functioning deficits and/or relative strengths found in MR (as well as unique patterns of neuroanatomical abnormalities) differ as a function of the etiological mechanism (see Pennington & Bennetto, 1998; Pulsifier, 1996). In other words, individuals with MR do not form a homogeneous group with respect to neuropsychological development or adaptive functioning (Loveland & Tunali-Kotoski, 1998).

Genetic/Organic MR

Evidence suggests that at least 50% of intelligence can be explained by heredity, and there are now over 1,000 different known organic causes of MR (Mash & Wolfe, 2002); however, the eight most common (in order of prevalence) are fetal alcohol syndrome, Down syndrome, fragile X, autism spectrum disorder, phenylketonuria, William syndrome, Prader–Willi syndrome, and Angelman syndrome. The first three causes together comprise approximately 33% of the total cases of MR with identifiable causes, and the scope of this chapter is therefore limited to a discussion of the neuropsychological profiles associated with these conditions. An excellent, comprehensive review of each of these causes and their neuropsychological sequalae can be found elsewhere in the literature (see Pulsifier, 1996). Following is a condensed summary of this work.

Fetal Alcohol Syndrome

Fetal alcohol syndrome (FAS) is reportedly the leading single cause of MR (see Pulsifier, 1996). The syndrome is caused and defined by maternal alcohol use during pregnancy and is characterized by stunted growth, unusual physical features (e.g., characteristic facies) and central nervous system (CNS) dysfunction that usually includes, but is not limited to, MR. Other CNS abnormalities that can occur include delay in reaching developmental milestones, microencephaly, muscular hypotonia, and hyperactivity. "Fetal alcohol effects" is the term used when full criteria for the syndrome are not met.

Associated neuropsychological deficits appear to be dose related. More specifically, the literature suggests that low doses of in uterine exposure can lead to slower mental and motor speed, behavioral impulsivity, and less complex play, while higher doses (i.e., ≥ two to three drinks per day) are associated with reduced fine motor coordination and sustained attention. Exposure to alcohol in the womb has also been shown to lead to abnormalities in language function (e.g., articulation disorders), certain aspects of executive function (i.e., abstract thinking, problem-solving, and judgment), grip strength, attention, arithmetic, and short-term memory. Overall IQ in children exposed to more than one-fifth ounces of alcohol per day *in utero* was found to be about 5 points below controls by age 4 in one study (see Mash & Wolfe, 2002). In general, Verbal IQ is also found to be generally below Performance IQ in this group.

In terms of adaptive functioning, individuals with FAS typically function well in daily living skills but are not as successful in demonstrating adequate social function (see Pulsifier, 1996). Further, this group is known for its persisting behavior difficulties that manifest in problems with feeding and sleeping, irritability, overactivity and distractability in school, difficulties with speech, social and emotional problems, and stereotypical behavior, such as head and body rocking or nail biting. Some of these individuals eventually meet criteria for conduct disorder, displaying antisocial behavior and later being at risk for alcohol abuse themselves. However, as a group, they are generally able to live independently with only limited support.

Down Syndrome

Affecting chromosome 21, Down syndrome (DS) is the most common genetic cause of MR, accounting for one-third of all moderate to severe MR in the world (Pulsifier, 1996). The syndrome is characterized by physical abnormalities including microcephaly, upward-slanting eyes, a broad neck, and hands that are small with in-curving fifth fingers. These individuals are also usually shorter in height than their non-DS counterparts, and obesity is common, as well as changes in hair, skin elasticity, motor skills, and skeletal structure that mimic an early aging process. In fact, some research suggests these individuals might be at higher risk for the later development of Alzheimer's disease.

Neuropsychologically, these individuals manifest severe language impairment, including deficits in articulation, phonology, and expressive syntax. In fact, their phrase structure may not exceed the 2nd-grade level in some cases. This is in contrast to their relatively preserved visuospatial abilities (Pezzini, Vicari, Volterra, Milani, & Ossella, 1999; Edgin, 2003). A related dissociation also occurs between visual and auditory sequential processing and memory, with specific sparing of the former; in particular, individuals with DS display difficulty recalling auditorily presented material. As in FAS, individuals with DS also display delayed motor abilities (e.g., with abnormally slower performance on Finger Tapping than comparison groups; Frith & Frith, 1974, and lack of expected hand asymmetry in dominant and nondominant hands upon finger tapping; Madhavan & Narayan, 1988), as well as motor clumsiness. They also do *not* show improvement on pursuit rotor tracking in contrast to the marked improvement seen in severe MR autistic groups and controls. These findings are consistent with neuroanatomical studies demonstrating that DS groups have smaller cerebellums compared to normals and relative to the rest of their own brains.

Adaptively, these individuals function higher than their cognitive abilities would suggest, with particular strengths in daily living and socialization skills relative to communication on the VABS (see Pulsifier, 1996). Further, they display higher than expected levels of social maturity/social adaptiveness given their mental age or IQ as compared to language-communication abilities. However, the relationship between adaptive skills and intelligence may differ with chronological age, with relative weakness in language and communication increasing with age but with eventual improvement in daily life skills (see Loveland & Tunali-Kotoski, 1998, for a review). Behaviorally, individuals with DS tend to function fairly well with low incidence of comorbid psychiatric illness. They are often able to live in group settings with limited support and are able to work in supported employment situations (Pulsifier, 1996).

Fragile X

Fragile X is the most common inherited cause of MR with a very distinct physical phenotype (at least in males) that emerges near puberty marked by macrocephaly, an elongated face, prominent jaw, long protruding ears, and macroorchidism (i.e., enlarged testicles) in adulthood. Similar changes occur in females as well, but to a much lesser extent. Males show developmental delays in sitting, walking, and speaking and are more often found to be hypotonic and temperamentally difficult. Intellectually, fragile X males usually present with moderate to severe MR, but a range is present, accompanied by primary neuropsychological deficits in attention, visual and verbal short-term memory, visual–spatial functioning (e.g., Block Design tends to be one of their lowest scores), and mental speed. Academic achievement tends to be better than expected given IQ level; however, math is a clear weakness. Males with fragile X demonstrate relative strengths in expressive and receptive vocabulary, as well as verbal comprehension; however, their speech is often marked by articulation problems and tends to be echolalic, perseverative, and dysfluent. In contrast to individuals with DS of similar intelligence, fragile X males display poorer conversational skills as well as abnormalities in the pragmatics of language (e.g., they will often avert their gaze when talking to another person; Pulsifier, 1996).

The adaptive profile of a male with fragile X on the VABS is one of relative strength in daily life skills, particularly personal and domestic skills. Relative weaknesses lie in socialization skills and communication abilities. Behaviorally, these individuals often engage in self-stimulatory behaviors, such as hand flapping, and self-injurious behaviors, such as head banging. They have also been found to be energetic, impulsive and inattentive, with most meeting criteria for attention-deficit/hyperactivity disorder (ADHD). A smaller subgroup of these individuals show autistic behavior, while others are anxious.

Females with fragile X present quite differently. First, only about half of these individuals have MR, and it is usually less severe than in males with the disorder (i.e., mild to moderate). While no initial developmental delays are evident, over time difficulties with attention, speech, and the pragmatics of language can emerge, and some of these females are diagnosed with learning disorders. However, many display no cognitive impairment at all, although typically there are relative weaknesses in visual–spatial skills and arithmetic, and behaviorally, these individuals tend to be shy and withdrawn with more frequent diagnoses of avoidant personality disorder or mood disorders when compared to non-fragile X females of similar IQ. Overall level of adaptive skills tends to be commensurate with the level of their intellectual development (Pulsifier, 1996).

Ideopathic/Cultural–Familial MR

Unlike the disorders mentioned earlier, there is no identifiable cause for MR in the majority of cases, especially milder forms (Mash & Wolfe, 2002). Ideopathic or cultural–familial MR is the "catchall" category used to describe lowered intelligence due to unidentified neurological and genetic disorders, as well as those cases of MR that are environmentally determined, collectively representing about one-half of all mild MR (see Pulsifier, 1996) and about 15–20% of MR in general (Mash & Wolfe, 2002). Ziegler and Hodapp (1986) posit subtypes, including the "familial retarded" in which at least one parent is also mentally retarded, the "polygenic isolates" in which parents are of average intelligence and provide generally nurturant environments, and finally, the "environmentally deprived" subtype in which poverty, parental neglect and/or psychopathology, or inadequate child care/educational experiences were experienced.

Physically, these individuals are frequently normal in appearance and health and have a mortality rate comparable to that of the general population. However, they may display mild physical abnormalities (e.g., abnormal skull shape). They are also usually from lower socioeconomic backgrounds and have relatives who are mentally retarded. Neuropsychologically, although there is some evidence to suggest these individuals display weakness in sequential as compared to simultaneous processing, slightly reduced achievement scores in arithmetic compared to reading, and better visual short-term memory compared to those with organic MR, there reportedly exists no distinct neuropsychological profile. Likewise, neuroanatomical variations found upon magnetic resonance imaging (MRI) scanning of children with idiopathic MR show only normal brain variations. Adaptively, these individuals generally live independently with minimal support and are able to integrate into the community. More than 80% find employment and 80% will marry (Pulsifier, 1996).

In sum, MR can occur in the context of several biological and/or environmental causes, each with a neuropsychological profile of relative strengths and weaknesses. Likewise, the organization of adaptive behavior also varies to some degree by etiology. Recall, for example, that while individuals with FAS, fragile X and DS all have relative strengths in daily living skills, FAS and fragile X individuals are much weaker in socialization skills as compared to the DS group. Thus, while some commonalities in MR can be discerned, there are also many differences. Further, there are numerous other conditions not reviewed here (e.g., 9p monosomy) that can be associated with MR and that have specific neuropsychological sequelae (see McSweeny, Wood, Chessare, & Kurczynski, 1993). Complicating matters even further is that the relationship between intellectual impairment and adaptive impairment varies by etiology (e.g., IQ is more strongly related to adaptive functioning in DS as compared to some of the other syndromes; Loveland, Tunali-Kotoski, 1998), and often individuals with mild MR (as documented by psychometric testing) are able to function well, all of which must be considered in the assessment of malingering in this population.

Given known prevalence rates, it is likely that the majority of cases seen in clinical practice are idiopathic in nature (and therefore cases of mild MR) or the result of FAS (and therefore mild to moderate MR). Thus, using the neuropsychological profiles associated with these two categories of MR to guide initial hypotheses about ability to pass effort tests is potentially the most productive course of action and would be summarized as follows:

Relative weaknesses	Relative strengths
Verbal IQ	Performance IQ
Auditory short-term memory	Visual short-term memory
Sequential processing	Simultaneous processing
Mental and physical speed	
Fine motor coordination	
Grip strength	
Simple and sustained attention	
Arithmetic	
Language function (e.g., articulation disorders)	
Abstract thinking, problem solving and judgment	

This profile raises the question as to whether an individual with MR would be able to pass standard effort test cutoffs. For example, given its reliance on sustained attention, would an individual with MR be able to pass the b Test? Likewise, given an MR individual's difficulty in arithmetic skills, would he/she be able to use multiplication necessary for successful passing of the Dot Counting Test? Further, while the field is moving in the direction of incorporating indices of response time to detect suboptimal effort in the general population (Victor & Abeles, 2004), is it likely that this method will be ineffective in population with MR given its members' slower mental and physical speeds? Would it be best to use only visual tests of effort given the relative strength in visual processing and memory in mentally retarded individuals?

To determine how practitioners are navigating the issue of effort testing in the differential diagnosis of actual versus feigned MR, we surveyed experts conducting these types of evaluations, the results of which are presented below.

SURVEY OF CURRENT PRACTICES IN FORENSIC PSYCHOLOGY

In January 2005, a confidential survey was sent to diplomates of the American Board of Forensic Psychologists (ABFP) to investigate current practices in detecting suspect effort in the context of forensic MR assessment. Of the 212 surveys distributed through a combination of electronic and land mail, 53 were completed and returned. However, three respondents indicated that they did not conduct medicolegal or criminal assessment and were therefore deemed inappropriate for the survey, rendering a final sample of 50.

Sample demographics are presented in Table 14.1. Notably, the sample was very sophisticated, with a mean of 23 years of experience in forensic work (consisting of an average 24% in medicolegal and 50% in criminal work). Further, 90% indicated they had testified or had their reports used in the context of a legal case over 100 times. Ninety percent also said they had been involved in a legal case in which they were asked to evaluate a litigant or criminal defendant for MR (constituting an average 10% of their total workload), with over half of the sample having experience with this 20 or more times in the context of forensic MR assessment.

Table 14.2 summarizes the current practice of respondents in relation to forensic MR assessment. The sample indicated that approximately 9% of the individuals assessed in the last year were faking MR. The most frequently used measures to assess feigned *intellectual* (vs. Axis I) deficits in MR evaluations were the following (in order of frequency): (1) Test of Memory Malingering (Tombaugh, 1996; 64%), (2) the Validity Indicator Profile (Frederick, 1997; 50%), (3) the Rey 15-Item Test (Rey, 1964, in Lezak, 2004; 44%), (4) the Structured Inventory of Reported Symptoms (Rogers, Bagby, & Dickens, 1992; 30%), (5) Word Memory Test (Green, Allen & Astner, 1996; Green & Astner, 1995; 14%), and finally, (6) the Dot Counting Test (Rey, 1941, in Lezak, 2004; Boone, Lu, et al., 2002; 10%). All but one respondent was aware of the Supreme Court's 2002 decision regarding execution of individuals with MR, and 96% of the sample felt that the problem of detecting feigned intellectual impairment was "very important" or "extremely important." Further, over half of the sample (i.e., 60%) felt "very confident" or "extremely confident" about their ability to accurately detect feigned MR. This is in spite of the paucity of research validating the use of these tests in samples with

TABLE 14.1. Demographics of Survey Respondents

Variable	Number	%	Mean	*SD*
Age			53.2	9.2
Gender				
Male	42	84%		
Female	8	16%		
Employment[a]				
Private practice	45	90%		
Academic	17	34%		
Jail/prison	4	6%		
Other	1	2%		
Retired	8	16%		
Geographic region				
Northeast	15	30%		
Southeast	19	38%		
Northwest	0	0%		
Southwest	8	16%		
Midwest	7	14%		
Did not respond		2%		
Forensic experience				
% medicolegal work		24%		
% criminal work		50%		
Number of years	23			
Number of times testified				
<10	0	0%		
10–50	2	4%		
51–99	3	6%		
>100	45	90%		
Experience with MR assessment				
Proportion of workload		10%		
Number of times testified				
<5	5	10%		
5–10	7	14%		
10–20	5	10%		
>20	33	66%		

Note. $n = 50$.
[a] Some respondents indicated more than one employment type.

TABLE 14.2. Practices of Survey Respondents

Variable	Number	%
Tests used to assess MR		
WAIS	48	96%
Other intelligence test	17	34%
Vineland		
ABS		
Other adaptive	11	22%
Projectives	1	2%
Use of collateral data	18	36%
% time use adaptive measures		57%
Tests used to assess feigned intellectual deficit		
TOMM	32	64%
VIP	25	50%
Rey 15-Item	22	44%
SIRS	15	30%
WMT	7	14%
Miller Forensic Assessment of Symptoms Test	6	12%
DCT	5	10%
Computerized Assessment of Response Bias	2	4%
VSVT	1	2%
Portland Digit Recognition Test	1	2%
MACT	1	2%
Importance of topic		
Not very important	0	0%
Somewhat important	1	2%
Very important	27	54%
Extremely important	21	42%
Confidence in detecting feigned MR		
Not at all confident	0	0%
Somewhat confident	16	32%
Very confident	24	48%
Extremely confident	6	12%
Depends on case	2	4%
Awareness of court decision	49	98%

Note. $n = 50$.

MR. Following is a review of how well these more popular tests stand up to empirical scrutiny.

FEIGNED MR: CAN WE ACCURATELY DETECT IT?

Mentally retarded and lower intelligence individuals (not to mention other important clinical samples) are typically excluded from effort test validation samples, which are instead normed on individuals with "normal" intelligence (Tombaugh, 2002; Weinborn et al., 2003). There is also surprisingly little in the general literature addressing the issue of using our effort measures with individuals of lowered intelligence. Our review uncovered only a handful of investigations examining the validity of widely used effort tests in this population.

Test of Memory Malingering

The suggestion that standard Test of Memory Malingering (TOMM; Tombaugh, 1996) cutoffs may be inappropriate for use with mentally retarded individuals has been made in the literature. For example, Weinborn et al. (2003) attempted to validate the TOMM in the context of a differential prevalence design examining factors that might have influenced subject performance (overall $N = 61$), including (but not limited to) MR (18% of sample) and borderline intellectual functioning (15% of sample). While the results provided general support for use of the TOMM, subjects with low intelligence were not segregated from those of normal intelligence. Analysis of the false positives (of which there were three) revealed they were all subjects with MR (two moderate and one mild). Given that there were 11 subjects with MR in total, 27% of subjects with MR were therefore misclassified by the TOMM, suggesting that the TOMM may not be appropriate for use with mentally retarded individuals. However, there were too few subjects to thoroughly valuate this hypothesis, and it was concluded by the authors that "the performance of individuals with mental retardation or borderline intellectual functioning on the TOMM has not been [adequately] evaluated" (Weinborn et al., 2003, p. 981).

Recently, Hurley and Deal (2006) examined the false-positive rate of the TOMM in a sample of subjects with known MR (i.e., they were living in residential facilities for individuals with MR and none of them had prior legal involvement; $N = 39$, mean age = 44.9, 96% Caucasian, Full Scale IQs ranging from 50 to 78) and found that standard TOMM cutoffs (< 45 on Trial 2) misclassified 41% of the sample. Similarly, Graue et al. (in press) compared a group of individuals with mild MR (with prior diagnosis, history of special education, currently in day-treatment programs, and *not* in litigation; $N = 26$) to demographically comparable community volunteers who were instructed to fake MR ($N = 25$), and they found that the TOMM misclassified 31% of their subjects with MR. However, Hurley and Deal (2006) noted that analyzing the change scores from Trial 1 to Trial 2 may provide important information, as about 85% of their sample with MR showed improvement.

Other studies have suggested that established TOMM cutoffs are appropriate for use with an MR sample. For example, Heinze and Purisch (2001) found the TOMM and other effort tests to be valid in a criminal population suspected of feigning incompetence to stand trial (many of whom were likely mentally retarded; $N = 57$); however, the authors suggested that not one of the tests alone was sufficiently sensitive. Unfortunately, there was no true MR comparison group so specificity, or the rate of false-positive identifications, could not be estimated. Preliminary data in a separate study using a sample of patients with known MR seen in the context of social security disability evaluation ($N = 19$; age range 18 to 60; WAIS-III Full Scale IQs range 57 to 70) indicated that the TOMM demonstrated 100% specificity (zero false positives) (Drwal, personal communication, October 2005).

Still other findings suggest that the TOMM cutoffs may be inappropriate only for use with samples of moderate (vs. mild) MR. Specifically, use of the TOMM was investigated with a sample of 60 psychiatric inpatients (mean age = 38, $SD = 10.9$, mean Full Scale IQ = 62.2, $SD = 5.1$; 68% male): (1) 29 psychiatric inpatients with MR (26 mild and 3 moderate); (2) two patients with MR without psychiatric illness (one mild, one moderate); and (3) 29 psychiatric patients of low intellectual functioning related to psychiatric disorder. Results indicated a moderate and significant correlation between Full Scale IQ and TOMM Trial 2 ($r = .55$, $p < .01$). In addition, there were only six failures (i.e., < 45 on Trial 2 and Retention), two of which were the individuals diagnosed with moderate MR. Given the calculated specificity of 90% for subjects with mild MR, the authors concluded that the TOMM was appropriate for use with these individuals but questioned its validity for use with individuals with moderate MR. Limitations of the study, however, include the confound of major mental illness in addition to the fact that some subjects were in litigation at the time of evaluation, including four of the six who failed the TOMM (Kennedy et al., 2005).

Validity Indicator Profile

Of note, there were no empirical data available for review of the Validity Indicator Profile (VIP) in a MR sample with the exception of that presented in the test manual, which provides data on 40 subjects with MR that were included in the validation sample. The test authors specifically state, "Most of the people (95%) with mental retardation were classified as invalid, supporting the caveat [that] the VIP test should not be administered to persons with obvious history of mental retardation" (Frederick, 1997, p. 8).

Rey 15-Item Test

In every investigation identified for this review, the validity of using standard cutoffs for the Rey 15-Item Test (Rey, 1964) with a MR population was called into question and a recent meta-analysis concluded that "one should use [the Rey-15] only with those who do *not* have mental retardation . . ." (Reznick, 2005, p. 542; italics added). The earliest study was that of Goldberg and Miller (1986) who administered the test to 50 acute psychiatric adult patients and 16 "intellectually deficient" adults (i.e., IQs ranging from 40 to 69 based on standard intelligence testing). While the psychiatric group all recalled at least 9 of the 15 items (and 92% recalled 11 or more), 37.5% of the mentally retarded group recalled 8 or fewer items and 0% recalled all 15. Citing the above-mentioned study in their review, Schretlen, Brandt, Krafft, and van Gorp (1991) noted that there was a paucity of literature examining the usefulness of the Rey 15-Item Test with individuals of low IQ, and they found a significant, modest correlation between total recall on the test and IQ in their own mixed sample of neurological, psychiatric and control subjects ($N = 193$; $r = .55$, $p < .001$). This is consistent with at least two other known investigations of IQ and Rey 15-Item Test performance ($N = 213$; $r = .53$, $p < .00$; Speigel, 2006; also see Hays, Emmons, & Lawson, 1993). Schretlen and his colleagues (2003) also observed that the number of repeated items on the tests correlated with IQ in the expected direction ($r = -.29$, $p < .001$). Further, although there was no strictly mentally retarded comparison group, the severe psychiatric group had the lowest IQ (mean = 73.6, $SD = 9.3$), and this group also performed the worst on the test, on the whole scoring below the standard cutoff (i.e., < 9 items).

Similarly, Hayes, Hale, and Gouvier (1997) investigated the utility of the Rey 15-Item Test (along with two other effort indicators) for predicting the malingering status of mentally retarded defendants ($N = 37$), including (1) 13 pretrial nonmalingerers with MR, (2) 18 subjects with MR found not guilty by reason of insanity (NGBRI), and (3) six malingerers with MR (as determined by treatment teams using stringent DSM-IV criteria and observation of inconsistent behavior). Results showed that the malingerers produced better memory scores than the groups with true MR, leading the authors of this paper to conclude that the Rey 15-Item Test should not be used among defendants with MR. Further, in a separate study of three groups of mentally retarded criminally insane men (9 malingerers, 12 nonmalingerers, and 18 dissimulators; mean IQ across groups = 62.4, mean $SD = 5.6$), when the Rey 15-Item Test was included in a discriminant function analysis along with two other effort indicators, the results were significant; however, only a little over half of the cases were correctly classified across groups, with unacceptable rates of specificity (i.e., over 25% of the nonmalingerers were incorrectly classified; Hayes, Hale, & Gouvier, 1998). Finally, a dissertation in progress examining Rey 15-Item performance (using standard cutoffs) in a sample

without identified incentive to feign cognitive impairment revealed specificity estimates of 69% and 77% for individuals with Full Scale IQs in the mentally retarded (i.e., less than 70; $N = 39$) and borderline (i.e., IQs between 70 and 79; $N = 49$) ranges of intellectual functioning, respectively. In addition, IQ subgroups (from borderline to high average) differed significantly (after controlling for the effects of education) on test performance; specifically, patients in the borderline range scored significantly lower than the low average ($p < .01$), average ($p < .01$) and high average ($p < .01$) groups (overall ratio $F = 14.48$; $N = 147$). Given that approximately one-third of the sample fell within the borderline or mentally retarded IQ ranges, the author suggested that the test was not appropriate (especially for exclusive use as an effort indicator) in the type of ethnically diverse, low socioeconomic status, urban population from which the sample was drawn (Speigel, 2006).

All the above-mentioned study designs were confounded by psychiatric and/or neurological impairment. There are only two studies to our knowledge that examined an uncontaminated sample of MR subjects. Marshall and Happe (in press) examined the performance of 100 patients with MR (WAIS-III Full Scale IQs ranged from 51 to 74), referred for standard clinical neuropsychological assessment, on five commonly used malingering tests, including the Rey 15-Item Test (cutoff < 9; Lezak, 2004) and the revised Rey 15-Item Test with recognition trial (cutoff < 20; Boone, Salazar, Lu, Warner-Chacon, & Razani, 2002). Subjects had a history of confirmed MR based on intelligence testing and a history of special education in school with no identified reason to feign cognitive impairment (i.e., they were already receiving support services from a county developmental disabilities division and had no reason to believe they were in danger of losing such services). Further, they were in good health with no drug, psychiatric, or medical history that might negatively affect their cognition. Results indicated that over half failed the recall portion of the test (mean recall = 7.5, $SD = 3.6$, $N = 69$) and over 80% failed the combined recall/recognition trial (mean combination score = 11.9, $SD = 7.2$, $N = 69$), further invalidating their use with this population and isolating the impact of IQ on effort test performance in a way that previous investigations had not. These results are consistent with the results of the Hurley and Deal (2006) study described earlier ($N = 39$ known subjects with MR), demonstrating that nearly 80% of their sample were misclassified by the Rey-15 Item Test using the < 9 cutoff.

Structured Interview of Reported Symptoms

In one of the same studies cited earlier, Hayes, Hale, and Gouvier (1998) tested three groups of men with MR from a state facility for the criminally insane ($N = 39$). Group 1 consisted of 12 pretrial nonmalingerers, Group 2 of 9 pretrial malingerers (as determined by an interdisplinary team), and

Group 3 consisted of 18 individuals faking good who were originally found to be NGBRI but desired release at the time of evaluation. Discriminant function analysis revealed that the Structured Interview of Reported Symptoms (SIRS; Rogers et al., 1992) alone produced a rate of 95% overall classification accuracy. However, methodological limitations included the fact that the sample was a mixed group with various comorbid psychiatric conditions, and group status was defined solely by an interdisciplinary team assessment of self-reported symptoms and behaviors. Further, the malingerers were likely feigning psychosis as well as MR, limiting the generalizability of these findings to settings in which only MR is being feigned.

In the Hurley and Deal (2006) sample of known individuals with MR, the SIRS misclassified 53.8% of the sample using the total score cutoff (> 76), 30.8% of the sample using one definite scale cutoff (≤ 1), and 30.8% of the sample using three probable scales cutoff (≥ 3). Curiously, however, the authors found no relationship between the SIRS total score and subject IQ (r = .03). The authors speculated that this finding may be a function of test structure, specifically that "yes" responses produce higher scores on the SIRS, and individuals with MR may be more likely to display this type of response bias (consistent with their tendency to make false confessions, recall Everington & Fulero, 1999).

Word Memory Test

The Word Memory Test (WMT; Green et al., 1996; Green & Astner, 1995) is the only effort test that includes in its manual validation data for an adult population with MR, specifically demonstrating that adults with Verbal IQs in the MR range (i.e., ≤ 70; mean Verbal IQ = 64) scored a mean of 96% (SD = 5) on the WMT, which is well above the standard cutoff of 85% (Green, 2003); however, specificity rates were not reported, and the authors specifically state that "future studies should determine WMT scores for motivated individuals of given VIQ ranges [as]. . . . Some adjustment to cut-off scores might be needed for people of very low VIQ" (Green et al., 1996, p. 13).

VIQ was also not found to significantly predict WMT performance in children and adults who had at least a third-grade reading level (Green & Flaro, 2003; Green, Lees-Haley, & Allen, 2002, respectively). In addition, children of lower intelligence reportedly passed a shorter version of the test, the Memory and Concentration Test (MACT; Green, 2004). More specifically, in a sample of 33 children, including 3 subjects with MR and 7 subjects with FAS, given the oral version of the MACT, all but one child passed the primary effort measures. However, the sample size used in this analysis was inadequate and the one false positive was a 17-year-old boy with a Full Scale IQ of 56, suggesting that while individuals with mild MR might be able to successfully complete the task, those bordering moderate levels of MR might have difficulty.

Dot Counting Test

The DOT Counting Test (DCT) combination E-score (Rey, 1941, in Lezak, 2004; Boone, Lu, et al., 2002) has proven to be inadequately specific when used with a MR sample seen for social security disability evaluations. Specifically, in a sample of 24 known MR subjects (based on prior intelligence testing) ranging in age from 18 to 60 (mean = 31.8; SD = 11.9) with WAIS-III Full Scale IQs ranging from 57 to 70 (mean = 65.4; SD = 3.6), there were 11 false positives using the E-score's generic cutoff of greater than or equal to 17, which corresponds to a specificity rate of only 54% (Drwal, personal communication, October 2005). Behavioral observations suggested that patients had particular difficulty applying multiplication skills to the task and instead required lengthy periods of time to individually count grouped dots. Limitations of this analysis include the fact that subjects had motive to feign; however, several indices were used to exclude malingerers, such as inconsistencies found between two IQ tests administered the same day (e.g., between the Kaufman Brief Intelligence Test [K-Bit; Kaufman & Kaufman, 1990] and the WAIS-III) or inconsistent reports (e.g., a patient stating he could not work secondary to depression but could find the energy to volunteer at church 40 hours per week); however, the exact number of those individuals excluded was not reported.

Likewise, use of the DCT was found inadequate in both defendants with MR and a sample of screened subjects with MR (i.e., without comorbid substance abuse, or psychiatric or neurological condition). Specifically, in one of the same studies cited previously, Hayes et al. (1997) investigated the utility of the DCT (along with the Rey 15-Item Test and another effort indicator) for predicting the malingering status of defendants with MR (N = 37) and found that the malingerers made fewer dot-counting errors than the other two groups with true MR, leading the authors of this paper to conclude that the DCT should not be used among defendants with MR. The DCT E-score (cutoff ≤ 17; Boone, Lu, et al., 2002) was also included in the Marshall and Happe (in press) study of effort test performance in a sample of 100 subjects with MR (WAIS-III Full Scale IQs ranged from 51 to 74). Results indicated that 70% failed (mean total score = 25.8, SD = 9.6, N = 69), invalidating use of standard DCT indices with this population.

However, use of only the total time to complete score from the DCT may be useful with a mentally retarded population. Hurley and Deal (2006) found that while their 39 known subjects with MR made errors on the test, they were (for the most part) able to complete a six-card version of the test in 180 or less seconds, a recommended cutoff from Paul, Franzen, Cohen, and Fremouw (1992) (i.e., only one subject exceeded this cutoff). Limitations of this study, however, includes the fact that it was conducted with a predominantly Caucasian sample, and that performance at different levels of IQ (mild, moderate, etc.) was not examined.

Other Free-Standing Effort Indicators

Malingering Scale

Schretlen and Arkowitz (1990) compared five different groups (N = 20 in each), including two groups of prison inmates instructed to fake insanity or MR and three comparison groups (psychiatric inpatients, adults with MR [mean IQ = 47, SD = 12] and prison inmate controls) on the Minnesota Multiphasic Personality Inventory (MMPI), the Bender Gestalt Test, and the Malingering Scale (MaS), a 90-item experimental measure partially adapted from existing intelligence scales and designed by the authors to measure malingered cognitive symptoms. While those instructed to fake MR obtained significantly lower scores on the MaS (mean = 46, SD = 21) than the inmate controls (mean = 84, SD = 4), the true MR group scored the worst (mean = 63, SD = 9), suggesting that this is not a test that true MR individuals can pass.

Victoria Symptom Validity Test

Loring, Lee, and Meador (2005) investigated the validity of established cut-off scores for the Victoria Symptom Validity Test (VSVT; Slick, Hopp & Strauss, 1995) in a clinical sample of epileptic patients and found that of the 15 subjects with low intellectual functioning (i.e., WAIS-III Full Scale IQs between 60 and 69), 9 subjects either scored in the normal range or perfectly on the test. These authors concluded that low intelligence will not by itself cause failure on the VSVT, although their data suggest an unacceptably low specificity level (i.e., 60%) in the subjects with MR as a group.

Digit Memory Test and the Letter Memory Test

Graue et al. (in press) also investigated the Digit Memory Test (DMT; Guilmette, Hart, Guiliano, & Leininger, 1994; cutoff < 90%) and the Letter Memory Test (LMT; Inman et al., 1998; cutoff < 93%) in their groups of individuals with mild MR and laboratory simulators and found that while the DMT produced close to acceptable levels of predictive accuracy (i.e., specificity = 85%; sensitivity = 76%), the LMT misclassified 42% of their mentally retarded sample.

Standard Cognitive Effort Indicators

Many investigators have recommended that standard cognitive tests (vs. free-standing indicators) be used to detect noncredible test performance to both decrease overall administration time and lessen the potential impact of coaching (e.g., Sherman, Boone, Lu, & Razani, 2002). What is known about the performance of individuals with MR on these indicators is summarized below.

Digit Span

Babikian, Boone, Lu, and Arnold (2006) investigated the use of the Digit Span (DS) age corrected scaled score (ACSS) and Reliable Digit Span (RDS; i.e., the sum of the longest string of digits repeated without error over two trials under both forward and backward conditions) for detecting suspect effort in three groups of patients: (1) a noncredible group (N = 66); (2) a credible mixed clinic group (N = 56); and (3) a control group (N = 32). Individuals with MR were included in the analyses. Using the recommended standard cutoff for ACSS (i.e., ≤ 5; found to have the best predictive accuracy by the authors in their sample), there were three false positives, all averaging a FSIQ approximately 18 points lower than patients above the cutoff (i.e., Full Scale IQ = 73.7 vs. 91.2). Similarly, the recommended cutoff for RDS (i.e., ≤ 6) yielded two false positives who had a mean Full Scale IQ approximately 15 points lower than patients above the cutoff (i.e., Full Scale IQ = 74.5 vs. 90.7). Collectively, these findings predict that use of the ACSS and RDS with individuals of borderline or MR-range IQ leads to unacceptable rates of false-positive error. Similarly, Marshall and Happe's (in press) analysis of 100 nonpsychiatric, healthy subjects with MR (described in more detail previously) found that 69% failed RDS (using a cutoff of ≤ 7; Greiffenstein, Baker, & Gola, 1994), and (even more striking) Graue et al. (in press) found that using the same RDS cutoff and a cutoff of ≤ 5 for ACSS misclassified 85% and 81% of their mentally retarded sample, respectively, providing fairly clear evidence to suggest DS scores are invalid for use in an MR population.

Finger Tapping Test

Arnold et al. (2005) investigated the use of the Finger Tapping Test (FTT) (three trials alternating between hands) in multiple clinical comparison groups, including a low IQ sample of 17 adults (i.e., IQ ≤ 70) with no history of neurological disorder. Results suggested standard cutoffs were inadequate and that there was a need for adjusted cutoffs with this group. Using a cutoff of ≤ 33 for dominant hand performance resulted in 89% specificity and 43% sensitivity in men with MR (mean = 41.5, SD = 6.3, N = 9), while dominant hand performance ≤ 28 was associated with 87% specificity and 61% sensitivity in women with MR (mean = 40.8, SD = 6.0, N = 8), suggesting that this effort indicator may show promise in the differentiation of feigned versus actual mental retardation. However, sensitivity was compromised at these cutoffs, and limitations of this study include the fact that dividing the sample by gender led to small subsamples that affected the stability of specificity values and therefore requires replication.

Wechsler Adult Intelligence Scale—Third Edition

Graue et al. (in press) investigated both the Wechsler Adult Intelligence Scale–Third Edition (WAIS-III) Vocabulary–Digit Span index (Iverson & Tulsky, 2003; Greve, Bianchini, Mathias, Houston, & Crouch, 2003; cutoff ≥ 4) and the Mittenberg Disciminant Function (Mittenberg et al., 2001; cutoff > .21) in their sample of subjects with mild MR and simulators. Results demonstrated that the simulators were able to produce overall WAIS-III scores comparable to the group with MR, and that neither of the aforementioned derived effort indicators were effective in discriminating the two groups. The latter misclassified 35% of the subjects with MR and the former, while associated with no false-positive identification, detected 0% of the laboratory simulators.

California Verbal Learning Test and Wechsler Memory Scale—Third Edition

Marshall and Happe (in press) included in their analysis the performance of their 100 mentally retarded subjects (see details of sample above) on the California Verbal Learning Test (CVLT) forced-choice recognition test (CVLT-FCR; Delis, Kramer, Kaplan, & Ober, 2000; cutoff < 3 correct; Millis, Putnam, Adams, & Ricker, 1995) and the Wechsler Memory Scale–Third Edition (WMS-III) Rarely Missed Items (RMI) index (WMS-III RMI cutoff ≤ 136; Killgore & DellaPietra, 2000). Notably, 89% of their subjects reportedly passed the CVLT-FCR and 91% passed the RMI using standard cutoffs, suggesting that they may be particularly promising for use with this population. However, the typical trouble with the former is that using significant below-chance cutoffs results in poor sensitivity; unfortunately, in the Marshall and Happe (in press) study there was no malingering comparison group in which to evaluate this measure of predictive accuracy. With regard to the latter, Killgore and DellaPietra (2000) report sensitivity and specificity estimates of 97% and 100%, respectively in their RMI validation sample of simulated malingerers ($N = 36$) and bona fide brain-damaged subjects ($N = 51$; 12% mentally retarded), suggesting that use of this index may provide protection against false-positive error while still preserving its ability to detect true feigning in this population. Additional support for use of the RMI was found by Leduc (2005) in a sample of individuals applying for Social Security Disability Insurance, many of whom are often mentally retarded.

MMPI-2

One article discussed the use of the MMPI-2 for identifying malingered MR. It concluded that the test would be inappropriate to use for this purpose for a number of reasons, including the fact that many individuals with MR never reach the eighth-grade reading (or oral comprehension) level required for assurance of understanding of test items. Keyes (2004) also points out that the test manual specifically states that the "usefulness of the information obtained by the MMPI-2 depends heavily on the ability of test subject to understand the test instructions, to comply with the requirements of the task, to comprehend and interpret the content of the items as they relate to him or her, and to record these self-attributions in a reliable way" (Hathaway & McKinley, 1989, p. 13). Keyes (2004) argues that an individual with MR would not have the level of concentration or verbal comprehension abilities required to complete the 567-item test. He further argues that the language of the test manual suggests results should never be used to provide the state with evidence that could or would result in execution of a criminal defendant. This is alarming, as the author points out, given that the Mississippi Supreme Court has apparently ruled that the MMPI-2 is required to rule out malingering in the assessment of inmates up for the death penalty (Keyes, 2004).

In sum, there is little research investigating the use of our standardly employed effort indicator cutoff scores in a population with MR, and much of research that is available is plagued with methodological limitations (inadequate sample sizes; contamination of samples; incentives to feign; lack of comparison groups, etc.) that limit the reliability and generalizability of findings. In spite of this, the results of the survey we conducted suggest that forensic neuropsychological assessments are conducted on a regular basis with this population, and of particular concern is the fact that over half of the survey sample respondents (i.e., 60%) felt "very confident" or "extremely confident" about their ability to accurately detect feigned MR. Further, our review of the available literature revealed unacceptable specificity rates (i.e., < 90%; Millis, 1992) for the use of some of the more commonly used effort indicators (i.e., VIP, Rey 15-Item Test, and DCT).

Preliminary Data from Our Lab

In an attempt to address some of these methodological limitations and to provide preliminary data concerning how our effort tests behave in a low IQ population, we present preliminary data examining the false-positive rates of several effort indicators in a selected sample of patients with (1) no motive to feign (i.e., they were not applying for disability compensation and not in litigation at the time of evaluation), (2) no history of dementia or schizophrenia, and (3) an IQ of less than or equal to 75 based on the WAIS-III (i.e., they met AAMR psychometric criteria for MR). Forty-two subjects referred to one of two outpatient mental health clinics in southern California for neuropsychological evaluation met these criteria and all completed a full neuropsychological testing battery that included the following effort indicators: Rey

15-Item Test (Rey, 1964, in Lezak, 2004) with recognition trial (Boone, Salazar, et al., 2002a), DCT (Rey, 1941, in Lezak, 2004; Boone, Lu, et al., 2002), Warrington Recognition Memory Test–Words (WRMT-W; Iverson & Franzen, 1994; Millis, 1992; note: administration was altered to include having subjects read each word aloud upon presentation prior to making a decision regarding pleasantness–whether they liked the word, did not like the word, or felt neutral about it), Rey Word Recognition Test (Greiffenstein, Gola, & Baker, 1995), RDS (Babikian et al., 2006; Greiffenstein et al., 1994), WAIS-III DS ACSS (Iverson & Franzen, 1994; Iverson & Tulsky, 2003; Suhr, Tranel, Wefel, & Barrash, 1997), Rey–Osterreith Effort Equation (Lu, Boone, Cozolino, & Mitchell, 2003), Rey–Osterreith (R-O)/Rey Auditory Verbal Learning Test (RAVLT) discriminant function (Sherman et al., 2002) ·d the FTT (Arnold et al., 2005).

Subjects ranged in age from 19 to 60 (mean = 40.6; *SD* = 10.3); 64% were female. Various ethnic groups were represented, including 36% Hispanics, 31% African Americans, 26% Caucasians, 2% Asian, 2% Middle Eastern, and 2% identifying themselves as "other." Number of years of education ranged from 4 to 16 (mean = 11.6; *SD* = 2.3). WAIS-III Full Scale IQ scores ranged from 56 to 75 (mean = 69.2; *SD* = 4.9), and thus all subjects fell into the classification of mild MR.

Table 14.3 provides means, standard deviation and range values for subjects' performance on each of the indicators. Displayed in Table 14.4 is the degree of specificity of each indicator (cutoffs used for each measures are also noted in the table). Defining "acceptable" levels of specificity as ≥ 90% (i.e., ≤ 10% false-positive rate), adequate values were found for only the WRMT-W (95%). Marginally acceptable levels of specificity were found for the FTT (87%); however, this is not surprising given that the newly recommended cutoffs for patients with MR were employed in this analysis (i.e., ≤ 28 for women, ≤ 33 for men; Arnold et al., 2005). Varying degrees of inad-ıte specificity (i.e., < 90) were found for the remainder of the indicators, including the R-O/RAVLT Discriminant Function (87%), DCT (76%), the R-O Effort Equation (74%), DS ACSS (62%), RDS (62%), the Rey 15-Item Test plus recognition (59%), and Rey Word Recognition (37%), with false-positive rates ranging from 13% up to 63%, overall suggesting the standardly employed cutoffs for these indicators are not appropriate to use when attempting to identify inadequate effort in individuals who may be mentally retarded. Adjusted cutoffs to meet appropriate levels of specificity are also presented in the table; however, the degree to which lowering the cutoffs affects these indicators' sensitivity for detecting malingering is unknown and is a topic addressed in more detail later. Finally, noteworthy is the fact that, examined by IQ band, subjects with Full Scale IQs falling in the 70s failed one to two indicators on average, while subjects in the 60–69 IQ group failed an average of three indicators, and those with IQs in the 50–59 range failed a mean of four indicators.

TABLE 14.3. Performance of Subjects on the Effort Indicators

Effort indicator	*n*	Mean	*SD*	Min.	Max.
Recognition Memory Test[a]	19	43.5	5.7	32	50
Finger Tapping Test[b]	30	41.6	10.3	22	72
R-O/RAVLT Discriminant Function[c]	40	0.75	1.06	-3.00	2.31
DCT E-score[d]	33	14.9	5.5	6.7	32.6
R-O Effort Equation[e]	34	49.9	12.2	23.5	68
DS ACSS[f]	42	6.4	2.1	3	13
RDS[g]	39	7	2	2	11
Rey-15 Item plus recognition[h]	34	20.8	6.2	9	30
Rey Word Recognition[i]	19	6.5	3.3	1	13

Note. Total n = 42.
[a] Recognition Memory Test = total number correct out of 50.
[b] FTT = average with dominant hand across three trials.
[c] R-O/RAVLT discriminant function = (.006 × RAVLT Trial 1) – (.062 × R-O delay) + (.354 × RAVLT recognition) – 2.508.
[d] DCT E-score = Mean ungrouped dot counting time + Mean grouped dot counting time + number of errors.
[e] R-O Effort = Copy + [(recognition true positives – atypical false positives) × 3].
[f] DS ACSS = WAIS-III DS ACSS.
[g] RDS = Total number of digits recalled forwards and backwards on both trials.
[h] Rey 15-Item Test plus recognition = [recall correct + (recognition correct – false-positive errors)].
[i] Rey Word Recognition = Total recognition – number of false positives.

An Additional Consideration: The Co-Occurrence of MR and Malingering

Our discussion thus far has rested on the assumption that actual versus feigned MR are mutually exclusive categories. However, we recently assessed a 41-year-old criminal defendant with a Full Scale IQ of 58, which was documented at age 18 prior to any legal difficulties. He was charged in the death of an infant and was tested at the request of his defense attorney regarding competency to stand trial. On Trial I of the TOMM the defendant repeatedly indicated that he could not recall which items he had been shown. Trial 1 was discontinued and he was shown the stimulus pictures a second time. On TOMM Trial 2 he obtained a score of 4 out of 50, a score well below chance and suggestive of his awareness of the correct answers despite his responses to the contrary. In addition, his presentation was noteworthy for exceedingly long pauses in reciting digits forward (an average of 13.5 seconds to recite four digits; cutoff > 6 seconds; Babikian et al., 2006), particularly low dominant-hand finger tapping speed (24.6; cutoff ≤ 33; Arnold et al., 2005), and inability to reliably identify letters or compute the most basic single-digit addition calculations despite attending school for 12 years. Thus, it is apparent that feigning can occur in some individuals with MR, and this is yet another complicating consideration when conducting these types of assessments.

TABLE 14.4. False-Positive Rate Associated with the Use of Effort Tests in a Population of True MR Negatives

Effort indicator	Cutoff	No. of false positives	n	Specificity[a]	Adjusted cutoff[b]	Specificity[a] adjusted cu.
Recognition Memory Test[c]	< 33	1	19	95%	—	—
FTT[d]	≤ 33 (men) ≤ 28 (women)	4	30	87%	≤ 32 ≤ 27	90%
R-O/RAVLT Discriminant Function[e]	≤ –.40	5	40	87%	≤ –.70	90%
DCT E-score[f]	≥ 17	8	33	76%	≥ 23	92%
R-O Effort Equation[g]	≤ 47	9	34	74%	≤ 29	92%
DS ACSS[h]	≤ 5	16	42	62%	≤ 3	93%
RDS[i]	≤ 6	15	39	62%	≤ 4	90%
Rey 15-Item plus recognition[j]	≤ 20	14	34	59%	≤ 11	91%
Rey Word Recognition[k]	≤ 5 (men) ≤ 7 (women)	12	19	37%	≤ 1 (men) ≤ 3 (women)	90%

Note. Total n = 42.
[a] Specificity = the proportion of true negatives who were correctly classified.
[b] Adjusted cutoffs are associated with ≥ .90 levels of specificity.
[c] Recognition Memory Test = total number correct out of 50.
[d] FTT = average with dominant hand across three trials.
[e] R-O/RAVLT discriminant function = (.006 × RAVLT Trial 1) – (.062 × R-O delay) + (.354 × RAVLT recognition) – 2.508.
[f] DCT E-score = Mean ungrouped dot counting time + Mean grouped dot counting time + number of errors.
[g] R-O Effort = Copy + [(recognition true positives – atypical false positives) × 3].
[h] DS ACSS = WAIS-III DS subtest ACSS.
[i] RDS = Total number of digits recalled forwards and backwards on both trials.
[j] Rey 15-Item Test plus recognition = recall correct + (recognition correct – false positive errors).
[k] Rey Word Recognition = Total recognition – number of false positives.

SUMMARY, INTEGRATION, AND DIRECTIONS FOR FUTURE RESEARCH

There is now even greater incentive for individuals to feign intellectual deficits in the context of forensic neuropsychological assessment. However, our ability to accurately detect, or rule out, this behavior is complicated by several factors, including 1) the fact that MR is caused by a number of different conditions that are both biologically and environmentally determined and appear to have specific neuropsychological sequelae associated with them; (2) that subjects with MR are typically excluded from effort test validation samples; (3) that the extent of the neuropsychological deficits would suggest that standard effort test cutoffs may not be appropriate for use with this population; and, finally, (4) that the extant empirical literature addressing this issue is limited but would suggest great caution be used in effort test interpretation as the likelihood of false-positive error is probably quite high; individuals of borderline and MR levels of intelligence can fail on average one to four effort tests in a standard battery even when putting forth their full effort. Of further concern is the fact that these tests are being routinely used in a population with MR for the purposes of assisting legal decision makers without adequate empirical support for such practice. Based on these findings as reviewed previously, we provide the following preliminary suggestions regarding clinical practice and future research.

Significantly below-chance performance on forced-choice effort measures would appear to provide definitive evidence of feigning, even in individuals with MR (e.g., the CVLT forced-choice recognition test; Marshall and Happe, in press). However, sensitivity rates (i.e., correct identification of actual malingerers) are poor at these levels, necessitating that additional effort indices be examined. The available data suggests that the WRMT-W may have particular promise for the differential diagnosis of actual versus feigned MR, with specificity rates (i.e., nonmalingerers correctly identified as such) above 90% when using standard cutoffs that have been associated with acceptable sensitivity values (Iverson & Franzen, 1994; Millis, 1992). The R-O/RAVLT discriminant function (Sherman et al., 2002) produced a specificity rate approaching 90% and may also be valuable in this type of assessment. Also promising is the use of the WMS-III RMI, which has been found to have high specificity in individuals with MR, and the index's creators (Killgore & DellaPietra, 2000) report sensitivity and specificity estimates of 97% and 100%, respectively in their RMI validation sample of simulated malingerers (N = 36) and bona fide brain-damaged subjects (N = 51; 12% mentally retarded), suggesting that use of this index may provide protection against false-positive error while still preserving its ability to detect true feigning in this population. Finally, the WMT and the DMT may be helpful in assessing effort in this population, but additional data are needed.

In contrast, the use of standard cutoffs for the SIRS, the DCT, the Rey 15-Item Test, RDS, DS ACSS, the FTT, Vocabulary-Digit Span, Mittenberg Discriminant Function, and even the TOMM may lead to an unacceptable rate of false-positive identifications within MR samples. Use of derived effort indicators from the WAIS-III appears to be particularly contraindicated, a not unexpected finding given that the issue at hand is one of intelligence. Although it might seem practical to establish cutoffs specific to this population, as pointed out by Marshall and Happe (in press), it appears that MR individuals score so low on some of these tests that simply establishing new cutoffs is inadequate. For example, in their study, lowering the cutoff for the Rey 15-Item Test to < 3 still incorrectly classified 17% of their mentally retarded sample. Similarly, the cutoff for the DCT had to be raised to ≥ 38 to ...ieve acceptable levels of specificity (i.e., > .90). Using these cutoffs would ...prove specificity, but this would occur at the expense of sensitivity. This was most clearly demonstrated by Graue et al. (in press); lowering the cutoff for TOMM Trial 2 and Retention to 30 raised specificity from 69% to 96% but lowered sensitivities to 56% and 60%, respectively. However, these same authors reported that a revised cutoff score of < 80% for the DMT produced sensitivity of 72%, specificity of 95%, and a hit rate of 88%. Further, lowering the cutoff to < 70% on the LMT resulted in sensitivity of 76%, specificity of 96%, and a hit rate of 88%. Thus, sensitivity was not necessarily unacceptably compromised on these tests; however, as the authors point out, these cut-scores require cross-validation as limitations of the study included small sample sizes, geographic restriction, and a simulation design.

It might also be useful to alter, or simplify, the tests given what we know about the ability of individuals with MR to successfully complete them. For example, it has been suggested that eliminating consideration of the time ratio between counting grouped and ungrouped dots on the DCT in individuals who may be mentally retarded is warranted given their difficulties employing multiplication strategies (Drwal, personal communication, October 2005). However, again this has a negative impact on sensitivity (i.e., consideration of errors alone, rather than a combination of errors and time, has been associated with lowered DCT sensitivity; Boone, Lu, et al., 2002).

It may also be useful to examine effort test scores in concert, rather than individually. Although it is fairly likely that an individual of lowered intelligence will fail one to four indicators, it is much less likely that such an individual will fail five or six according to our preliminary data. Hayes et al. (1998) found that the use of three effort indicators (the DCT, the Rey 15-Item Test, and the M Test) correctly classified 100% of their sample; however, the false-positive rates based on failure on at least one of the four effort indices (i.e., the SIRS, the DCT, the Rey 15-Item Test, and the M Test) ranged from 13% to 27%. Likewise, Graue et al. (in press) found that 69% of their participants with MR failed at least one of the multiple effort tests administered in their study. Discriminant function analyses may be useful in identifying weighted combinations of test results specific to individuals with actual MR.

Qualitative analysis of the types of responses or errors made by patients may also hold promise. For example, in the study by Marshall and Happe (in press), it was very rare for any of their subjects with MR to make even one dyslexic, false-positive recognition error on the Rey 15-Item plus recognition test. Further, there were some data to suggest that individuals with MR show a "yes" response bias (e.g., on Logical Memory recognition), that is not likely to be displayed by malingerers (Marshall & Happe, in press). In fact, it is this response bias that the authors believe underlies the ability of individuals with MR to pass the Logical Memory RMI of the WMS-III ("yes" is the correct answer to all six questions that comprise this test). Recall this was a speculation made by Hurley and Deal (2006) with regard to the high false-positive rate of subjects with MR on the SIRS. Development of tests that capitalize on unique response biases found in samples with MR may be a particularly promising area of future research. In addition, more creative and innovative paradigms are needed.

As in any case in which the individual is suspected of not putting forth adequate effort, observations of behavioral inconsistencies can be useful. It has been suggested that malingerers might hold inaccurate beliefs about the abilities of mentally handicapped, focusing only on achieving a low IQ with little consideration of other aspects of their presentation such as their behavioral, conversational, or social skills presentation (Johnstone & Crooke, 2003). Asking yourself, "Is the patient able to understand verbal instructions?" for example, or making note of the linguistical structure of their speech and any indication of an articulation disorder would also be informative. Listening and making note of language used during conversation can be a good indicator of whether a person's true ability exceeds test performance (Morrison, 1994). Further, it is always important to explore several sources of information (i.e., test scores, clinical interview, and collateral information from original judge, defense attorney, court records, and educational records) given that individuals with true MR are likely to have documentation of their mental functioning over time (Miller & Germain, 1988). Access to this information will likely enhance one's ability to identify feigning, although these data will never eclipse the need for objective effort measures specifically validated in a population with MR.

Ultimately, it may be that different effort tests are required for a population with MR, rather than simply importing for use tests developed on individuals of normal intelligence. It is likely that particular effort tests are effective in some IQ ranges and not others. Future research in this area is needed to examine the utility of primarily visual tests with use of unconfounded samples with MR (who have no incentive to feign) and appropriate comparison groups.

REFERENCES

American Association on Mental Retardation. (1992). *Mental retardation: Definitions, classification and systems of support* (9th ed.). Washington, DC: Author.

American Association on Mental Retardation. (2002). *Mental retardation: Definitions, classification and systems of support* (10th ed.). Washington, DC: Author.

American Bar Association. (1984). *Criminal Justice Mental Health Standards*. Retrieved August 31, 2005, from *www.abanet.org/crimjust/standards/mentalhealth_toc.html*.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

Arnold, G., Boone, K., Dean, A., Wen, J., Nitch, S., Lu, P., et al. (2005). Sensitivity and specificity of finger tapping test scores for the detection of suspect effort. *The Clinical Neuropsychologist, 19*, 105–120.

A-kins v. Virginia, 536 U.S. 304, 153 L.Ed.2d 335 (2002).

ikian, T., Boone, K. B., Lu, P., & Arnold, G. (2006). Sensitivity and specificity of various digit span scores in the detection of suspect effort. *The Clinical Neuropsychologist, 20*, 145–159.

Black, A. H., & Davis, L. J. (1966). The relationship between intelligence and sensorimotor proficiency in retardates. *American Journal of Mental Deficiency, 71*, 55–59.

Boone, K. B., Lu, P., Back, C., King, C., Lee, A., Philpott, L., et al. (2002). Sensitivity and specificity of the Rey Dot Counting Test in patients with suspect effort and various clinical samples. *Archives of Clinical Neuropsychology, 17*, 625–642.

Boone, K. B., Salazar, X., Lu, P., Warner-Chacon, K., & Razani, J. (2002). The Rey 15 item Recognition Trial: A technique to enhance sensitivity of the Rey 15-Item Memorization Test. *Journal of Clinical and Experimental Neuropsychology, 24*, 561–573.

Brodsky, S. L., & Galloway, V. A. (2003). Ethical and professional demands for forensic mental health professionals in post-Atkins era. *Ethics and Behavior, 13*, 3–9.

Bruininks, R. H., McGrew, K., & Maruyama, G. (1988). Structure of adaptive behavior in samples with and without mental retardation. *American Journal on Mental Retardation, 93*, 265–272.

iinks, R. H., Thurlow, M., & Gilman, C. J. (1987). Adaptive behavioral and mental retardation. *The Journal of Special Education, 21*, 69–88.

Calvert, E. J., & Crozier, W. R. (1978). An analysis of verbal–performance intelligence quotient discrepancies in the Wechsler Adult Intelligence Scale results of mentally subnormal hospital patients. *Journal of Mental Deficiency Research, 22*, 147–153.

Davis, L. A. (2000). *People with mental retardation in the criminal justice system*. Retrieved July 27, 2006, from *www.thearc.org/faqs/crimjustice.doc*.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000). California Verbal Learning Test (2nd ed.). San Antonio, TX: Psychological Corporation.

DeVault, S., & Long, D. (1988). Adaptive behavior, malingering and competence to waive right: A case study. *American Journal of Forensic Psychology, 6*, 3–15.

Dolán, M. (2005, February 11). Inmates can appeal based on their IQ: State justices define mental disability, giving dozens on death row a chance for life terms. *The Los Angeles Times*, pp. A1, A31.

Edgin, J. O. (2003). A neuropsychological model for the development of the cognitive profiles in mental retardation syndromes: Evidence from Down syndrome and Williams syndrome. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 64*, 1522.

Ellis, N. R. (1963). The stimulus trace and behavioral inadequacy. In N. R. Ellis (Ed.), *Handbook of mental deficiency* (pp. 134–158). New York: McGraw-Hill.

Everington, C., & Fulero, S. (1999). Competence to confess: Measuring understanding and suggestibility of defendants with mental retardation. *Mental Retardation, 37*, 212–220.

Everington, C., & Keyes, D. W. (1999). Mental retardation: Diagnosing mental retardation in criminal proceedings: The critical importance of documenting adaptive behavior. *Forensic Examiner, 8*, 31–34.

Fogelmann, C. J. (Ed.). (1975). *AAMD Adaptive Behavior Scale Manual*. Washington, DC: American Association on Mental Deficiency.

Frederick, R. I. (1997). *Validity Indicator Profile manual*. Minneapolis, MN: National Computer Services.

Frederick, R. I. (2000a). Mixed group validation: A method to address the limitations of criterion group validation in research on malingering detection. *Behavioral Sciences and the Law, 18*, 693–718.

Frederick, R. I. (2000b). A personal floor effect strategy to evaluate the validity of performance on memory tests. *Journal of Clinical and Experimental Neuropsychology, 22*, 720–730.

Frith, U., & Frith, C. D. (1974). Specific motor disabilities in Down's Syndrome. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 15*, 293–301.

Goldberg, J. O., & Miller, H. R. (1986). Performance of psychiatric inpatients and intellectually deficient individuals on a task that assesses the validity of memory complaints. *Journal of Clinical Psychology, 42*, 797–795.

Graue, L. O., Berry, D. T. R., Clark, J. A., Sollman, M. J., Cardi, M., Hopkins, J., et al. (in press). Identification of feigned mental retardation using the new generation of malingering detection instruments: Preliminary findings. *The Clinical Neuropsychologist*.

Green, P. (2003). *Green's Word Memory Test (WMT) for Windows User's Manual* [Online]. Edmonton, Alberta, Canada: Green's Publishing. Available at *www.wordmemorytest.com*.

Green, P. (2004). *Green's Memory and Concentration Test (MACT) for Windows User's Manual* [Online]. Edmonton, Alberta, Canada: Green's Publishing. Available at: *www.wordmemorytest.com*.

Green, P., Allen, L., & Astner, K. (1996). *Manual for Computerised Word Memory Test, U.S. version 1.0*. Durham, NC: Cognisyst.

Green, P., & Astner, K. (1995). *The Word Memory Test*. Edmonton, Alberta, Canada: Neurobehavioural Associates.

Green, P., & Flaro, L. (2003). Word Memory Test performance in children. *Child Neuropsychology, 9*, 189–207.

Green, P., Lees-Haley, P., & Allen, L. (2002). The Word Memory Test and the validity of neuropsychological test scores. *Forensic Neuropsychology, 2*, 97–124.

Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment, 6*, 218–224.

Greiffenstein, M. F., Gola, T., & Baker, W. J. (1995). MMPI-2 validity scales versus domain specific measures in detection of factitious traumatic brain injury. *Clinical Neuropsychologist, 9*, 230–240.

Greve, K. W., Bianchini, K. J., Mathias, C. W., Houston, R. J., & Crouch, J. A. (2003). Detecting malingered performance on the Wechsler Adult Intelligence Scale: Validation of Mittenberg's approach in traumatic brain injury. *Archives of Clinical Neuropsychology, 18*, 245–260.

Grossman, H. J. (Ed.). (1983). *Classification in mental retardation.* Washington, DC: American Association on Mental Deficiency.

Gudjonsson, G. H. (2002). *The psychology of interrogations and confessions: A handbook.* Chichester, UK: Wiley.

Guilmette, T. J., Hart, K. J., Guiliano, A. J., & Leininger, B. E. (1994). Detecting simulated memory impairment: Comparison of the Rey Fifteen-Item Test and the Hiscock Forced-Choice Procedure. *The Clinical Neuropsychology, 8*, 283–294.

.athaway, S. R., & McKinley, J. C. (1989). *Minnesota Multiphasic Personality Inventory–2.* Minneapolis: University of Minnesota Press.

Hawkins, G. D., & Cooper, D. H. (1990). Adaptive behavioral measures in mental retardation research: Subject description in AJMD/AJMR articles. *American Journal on Mental Retardation, 94*, 654–660.

Hawthorne, X. (2005, February 4). State high court to rule on mental retardation case: Death row inmates seeking IQ review. *Los Angeles Times*, pp. A1, A24.

Hayes, J. S., Hale, D. B., & Gouvier, W. D. (1997). Do tests predict malingering in defendants with mental retardation? *Journal of Psychology, 131*, 575–576.

Hayes, J. S., Hale, D. B., & Gouvier, W. D. (1998). Malingering detection in a mentally retarded forensic population. *Applied Neuropsychology, 5*, 33–36.

Hays, J. R., Emmons, J., & Lawson, K. A. (1993). Psychiatric norms or the Rey 15-Item visual memory test. *Perceptual and Motor Skills, 76*, 1331–1334.

Heinze, M. C., & Purisch, A. D. (2001). Beneath the mask: Use of psychological tests to detect and subtype malingering in criminal defendants. *Journal of Forensic Psychology Practice, 1*, 23–52.

Huberty, T. J., Koller, J. R., & Ten Brink, T. D. (1980). Adaptive behavior in the definition of mental retardation. *Exceptional Children, 46*, 258–261.

.rley, K. E., & Deal, W. P. (2006). Assessment instruments measuring malingering used with individuals who have mental retardation: Potential problems and issues. *Mental Retardation, 44*, 112–119.

Iki, H., & Kusano, K. (1986). Muscle strength training of mentally retarded children. *Japanese Journal of Special Education, 24*, 20–26.

Inman, T. H., Vickery, C. D., Berry, D. T. R., Lamb, D., Edwards, C., & Smith, G. T. (1998). Development and initial validation of a new procedure for evaluating adequacy of effort during neuropsychological testing: The Letter Memory Test. *Psychological Assessment, 10*, 128–139.

Iverson, G. L., & Franzen, M. D. (1994). The Recognition Memory Test, Digit Span, and Knox Cube Test as markers of malingered memory impairment. *Assessment, 1*, 323–334.

Iverson, G. L., & Tulsky, D. S. (2003). Detecting malingering on the WAIS-III unusual Digit Span performance patterns in the normal population and in clinical groups. *Archives of Clinical Psychology, 18*, 1–9.

Jackson v. Indiana, 406 U.S. 715, 32 L.Ed.2d 435 (1972).

Johnstone, L., & Cooke, D. J. (2003). Feigned intellectual deficits on the Wechsler Adult Intelligence Scale—Revised. *British Journal of Clinical Psychology, 42*, 303–318.

Kaufman, A. S., & Kaufman, N. L. (1990). *Kaufman Brief Intelligence Test.* Minneapolis, MN: NCS.

Kennedy, C. Shaver, S., Weinborn, M., Manley, J., Broshek, D., & Marcopulos, B. (2005, November). *Use of the Test of Memory Malingering (TOMM) in Individuals with FSIQ below 70.* Poster session presented at the 25th annual meeting of the National Academy of Neuropsychology (NAN), Tampa, FL.

Keyes, D. W. (2004). Use of the Minnesota Multiphasic Personality Inventory (MMPI) to identify malingering mental retardation. *Mental Retardation, 42*, 151–153.

Keyes, D., Edwards, W., & Perske, R. (1997). People with mental retardation are dying—Legally. *Mental Retardation, 35*, 59–63.

Killgore, W., & DellaPietra, L. (2000). Using the WMS-III to detect malingering: Empirical validation of the rarely missed index. *Journal of Clinical and Experimental Neuropsychology, 22*, 761–771.

Lanzkron, J. (1963). Murder and insanity: A survey. *American Journal of Psychiatry, 119*, 754–758.

Leduc, L. G. (2005). Detection of dissimulation among social security disability applicants using the WMS-III. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 66*, 1723.

Lezak, M. (2004). *Neuropsychological Assessment* (4th ed.). New York: Oxford University Press.

Lin, Z., & Zhang, Z. (2002). Study on the speed of hand tapping in mentally retarded children [Abstract]. *Chinese Journal of Clinical Psychology, 10*, 230–231.

Lipkin, P. (1991). Epidemiology of the developmental disabilities. In A. J. Caputec & P. K. Accardo (Eds.), *Developmental disabilities in infancy and childhood* (pp. 43–63). Baltimore: Brookes.

Loring, D. W., Lee, G. P., & Meador, K. J. (2005). Victoria Symptom Validity Test performance in non-litigating epilepsy surgery candidates. *Journal of Clinical and Experimental Neuropsychology, 27*, 610–617.

Loveland, K. A., & Tunali-Kotoski, B. (1998). Development of adaptive behavior in persons with mental retardation. In R. M. Hodapp & J. A. Burack (Eds.), *Handbook of mental retardation and development* (pp. 521–541). New York: Cambridge University Press.

Lu, P. H., Boone, K. B., Cozolino, L., & Mitchell, C. (2003). Effectiveness of the Rey Osterreith Complex Figure Test and the Meyers and Meyers Recognition Trial in the detection of suspect effort. *The Clinical Neuropsychologist, 17*, 426–440.

Luckasson, R., Coulter, D. L., Polloway, E. A., Reiss, S., Schalock, R. L., Snell, M. E., et al. (1992). *Mental retardation: Definition, classification and systems of support.* Washington, DC: American Association on Mental Retardation.

MacKenzie, S., & Hulme, C. (1987). Memory span development in Down's syndrome severely subnormal and normal subjects. *Cognitive Neuropsychology, 4*, 303–319.

Madhaven, T., & Narayan, J. (1988). Manual asymmetries in persons with Down's syndrome. *Indian Journal of Psychiatry, 30*, 193–195.

Marshall, P. S., & Happe, M. (in press). The performance of individuals with mental retardation on cognitive tests assessing effort and motivation. *The Clinical Neuropsychologist.*

Mash, E. J., & Wolfe, D. A. (2002). *Abnormal child psychology* (2nd ed.). Belmont, CA: Wadsworth.

McCaffrey, R. J., & Isaac, W. (1985). Preliminary data on the presence of neuropsychological deficits in adults who are mentally retarded. *Mental Retardation, 23*, 63–66.

McLaren, J., & Bryson, S. E. (1987). Review of recent epidemiological studies of mental retardation: Prevalence, associated disorders and etiology. *American Journal of Mental Retardation, 92*, 243–254.

McSweeny, A. J., Wood, G. L., Chessare, J. B., & Kurczynski, T. W. (1993). Cognitive and language anomalies in two brothers with a rare form of genetic abnormality (deletion of 9p24). *Clinical Neuropsychologist, 7*, 460–466.

Miller, R. D., & Germain E. J. (1988). The retrospective evaluation of competency to stand trial. *International Journal of Law and Psychiatry, 11*, 113–125.

Millis, S. R. (1992). The recognition memory test in the detection of malingered and exaggerate memory deficits. *Clinical Neuropsychologist, 6*, 406–414.

Millis, S. R., Putnam, S. H., Adams, K. N., & Ricker, J. H. (1995). California Verbal Learning Test in the detection of incomplete effort in neuropsychological evaluation. *Psychological Assessment, 7*, 463–471.

Mittenberg, W., Theroux, S., Aguila-Puentas, G., Bianchini, K. J., Greve, K. W., & Rayls, K. (2001). Identification of malingered head injury on the Wechsler Adult Intelligence Scale–3rd Edition. *The Clinical Neuropsychologist, 15*, 440–445.

Morrison, M. W. (1994). The use of psychological tests to detect malingered intellectual impairment. *American Journal of Forensic Psychology, 12*, 47–54.

Nettelbeck, T. (1986). Inspection time as mental speed in mildly mentally retarded adults: Analysis of eye gaze, eye movement and orientation. *American Journal of Mental Deficiency, 91*, 78–91.

Nihira, K., Foster, R., Shellhaas, M., & Leland, H. (1974). *AAMD Adaptive Behavior Scale Manual.* Washington, DC: American Association on Mental Deficiency.

Nihira, K., Leland, H., & Lambert, N. (1993). *AAMR Adaptive Behavior Scale–Residential and Community* (2nd ed.). Austin, TX: Pro-Ed.

Paul, D., Franzen, M. D., Cohen, S. H., & Fremouw, W. (1992). An investigation into the reliability and validity of two tests used in the detection of dissimulation. *International Journal of Neuropsychology, 14*, 1–9.

Pennington, B. F., & Bennetto, L. (1998). Toward a neuropsychology of mental retardation. In J. A. Burack, R. M. Hodapp, & E. Zigler (Eds.), *Handbook of mental retardation and development* (pp. 80–114). New York: Cambridge University Press.

Pezzini, G., Vicari, S., Volterra, V., Milani, L., & Ossella, M. T. (1999). Children with Williams Syndrome: Is there a single neuropsychological profile? *Developmental Neuropsychology, 15*, 141–155.

Pulsifier, M. B. (1996). The neuropsychology of mental retardation. *Journal of International Neuropsychological Society, 2*, 159–176.

Rey, A. (1964). *The clinical exam in psychology.* Paris: Presses Universitaries de France.

Reznick, L. (2005). The Rey 15-Item Memory Test for malingering: A meta-analysis. *Brain Injury, 19*, 539–543.

Rogers, R., Bagby, R. M., & Dickens, S. E. (1992). *Structured Interview of Reported Symptoms.* Odessa, FL: Psychological Assessment Resources.

Roszkowski, M. J., & Snelbecker, G. E. (1981). Configuration of WAIS subtest scores in mental retardation: Further evidence. *Psychological Reports, 48*, 1006.

Sattler, J. (1988). *Assessment of children* (3rd ed.). San Diego, CA: Sattler.

Schlesinger, L. B. (2003). A case study involving competency to stand trial: Incompetent defendant, incompetent examiner, or "malingering by proxy"? *Psychology, Public Policy and Law, 9*, 381–399.

Schretlen, D., & Arkowitz, H. (1990). A psychological test battery to detect prison inmates who fake insanity or mental retardation. *Behavioral Sciences and the Law, 8*, 75–84.

Schretlen, D., Brandt, J., Krafft, L., & van Gorp, W. (1991). Some caveats in using the Rey 15-Item Memory Test to detect malingered amnesia. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3*, 667–672.

Sherman, D. S., Boone, K., Lu, P., & Razani, J. (2002). Re-examination of the Rey Auditory Verbal Learning Test/Rey Complex Figure discriminant function to detect suspect effort. *The Clinical Neuropsychologist, 16*, 242–250.

Slick, D., Hopp, G., & Strauss, E. (1995). Victoria Symptom Validity Test. Odessa, FL: Psychological Assessment Resources.

Sparrow, S. S., Balla, D., & Cicchetti, D. V. (1984). *Vineland Adaptive Behavior Scales* (Survey ed.). Circle Pines, MN: American Guidance Service.

Speigel, E. (2006). *The Rey 15-Item Memorization Test: Performance in under-represented populations in the absence of obvious motivation to feign neurocognitive symptoms.* Unpublished doctoral dissertation, Fuller Theological Seminary, Pasadena, CA.

Sprague, R. L., & Quay, H. C. (1966). A factor analytic study of the responses of mental retardates. *American Journal of Mental Deficiency, 70*, 595–600.

Spreen, O., & Strauss, E. (1998). *A compendium of neuropsychological tests: Administration, norms and commentary* (2nd ed.). New York: Oxford University Press.

Stevens, K. B., & Randall, J. (2006). Adaptive behavior, mental retardation and the death penalty. *Journal of Forensic Psychology Practice, 6*, 1–29.

Suhr, J., Tranel, D., Wefel, J., & Barrash, J. (1997). Memory performance after head injury: Contributions of malingering, litigation status, psychological factors, and medication use. *Journal of Clinical and Experimental Neuropsychology, 19*, 500–514.

Tombaugh, T. N. (1996). *Test of Memory Malingering.* Toronto, Ontario, Canada: Multi-Health Systems.

Tombaugh, T. N. (2002). The Test of Memory Malingering (TOMM) in forensic psychology. *Journal of Forensic Neuropsychology, 2*, 69–96.

Victor, T., & Abeles, N. (2004). Coaching clients to take psychological and neuropsychological tests: A clash of ethical obligations. *Professional Psychology: Research and Practice, 35*, 373–379.

Weinborn, M., Orr, T., Woods, S. P., Conover, E., & Feix, J. (2003). A validation of the Test of Memory Malingering in a forensic psychiatric setting. *Journal of Clinical and Experimental Neuropsychology, 25*, 979–990.

Ziegler, E. & Hodapp, R. M. (1986). *Understanding mental retardation.* Boston: Cambridge University Press.